

## CONTENTS

<b>6 Mixed Model Analysis</b>	<b>2</b>
6.1 Normal-theory mixed model . . . . .	4
6.2 Analysis of Mixed Linear Models . . . . .	20
6.3 Generalized Least Squares (GLS) Estimation . . . . .	21
6.4 Variance component estimation . . . . .	23
6.4.1 Rules for Expected Mean Squares . . . . .	23
6.4.2 ANOVA method (Method of Moments)	30
6.4.3 Properties of ANOVA methods for variance component estimation . . . . .	33
6.5 Estimation of fixed effects . . . . .	43
6.6 Prediction of random effects . . . . .	43

## 6 Mixed Model Analysis

Basic model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where

$\mathbf{X}$  is a  $n \times p$  model matrix of known constants

$\boldsymbol{\beta}$  is a  $p \times 1$  vector of “fixed” unknown parameter values

$\mathbf{Z}$  is a  $n \times q$  model matrix of known constants

$\mathbf{u}$  is a  $q \times 1$  random vector

$\mathbf{e}$  is a  $n \times 1$  vector of random errors

with

$$E(\mathbf{e}) = \mathbf{0} \quad V(\mathbf{e}) = \mathbf{R}$$

$$E(\mathbf{u}) = \mathbf{0} \quad V(\mathbf{u}) = \mathbf{M}$$

$$Cov(\mathbf{e}, \mathbf{u}) = 0$$

Then

$$\begin{aligned} E(\mathbf{y}) &= \\ V(\mathbf{y}) &= \end{aligned}$$

### 6.1 Normal-theory mixed model

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}\right)$$

Then,  $\mathbf{y} \sim$

**Example 1.** Random Blocks

Comparison of four processes for producing penicillin

$$\left. \begin{array}{l} \text{Process } A \\ \text{Process } B \\ \text{Process } C \\ \text{Process } D \end{array} \right\} \quad \begin{array}{l} \text{Levels of a "fixed"} \\ \text{treatment factor} \end{array}$$

Blocks correspond to different batches of an important raw material, corn steep liquor

- Random sample of five batches
- Split each batch into four parts:

- run each process on one part
- randomize the order in which the processes are run within each batch

Here, batch effects are considered as *random* block effects.

- Batches are sampled from a population of many possible batches

- To repeat this experiment you would need to use a different set of batches of raw material

**Model:**

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

Yield  
for the  
i-th process  
applied  
to the  
j-th batch  
where  
averaging  
across the  
entire  
population  
of  
possible  
batches

$$\begin{array}{ll} \beta_j \sim NID(0, \sigma_{\beta}^2) \\ e_{ij} \sim NID(0, \sigma_e^2) \end{array}$$

and any  $e_{ij}$  is independent of any  $\beta_j$ . Here  $\mu_i =$

Restrict model with  $\alpha_4 = 0$ . Then

- $\mu =$
- $\alpha_i =$

In R we could use the "treatment" constraints where  $\alpha_1 = 0$ . Then

- $\mu =$
- $\alpha_i =$

Alternatively, we could choose the solution to the normal equations given by "sum" constraints.

- $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 =$
- $\mu =$
- $\alpha_i =$
- Correlation among yields for runs on the same batch:  $\rho =$

Variance-covariance structure:

- $V(Y_{ij}) =$

In R we could use the "treatment" constraints where  $\alpha_1 = 0$ . Then

- Different runs on the same batch:  $Cov(Y_{ij}, Y_{kj}) =$

- Results for runs on different batches are uncorrelated (independent):

$$Cov(Y_{ij}, Y_{k\ell}) = 0 \quad \text{for } j \neq \ell$$

- Results from the four runs on a single batch:

This special type of covariance structure is called *compound symmetry*

- Write this model as  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$

Here  
 $\mathbf{M} =$   
 $\mathbf{R} =$

**Example 2.** Hierarchical Random Effects Model  
 Analysis of sources of variation in a process used to monitor  
 the production of a pigment paste.  
**Current Procedure:**

- Sample barrels of pigment paste
- One sample from each barrel
- Send the sample to a lab for determination of moisture content

**Measured Response:** ( $Y$ ) moisture content of the pigment paste (units of one tenth of 1%).

**Problem:** Variation in moisture content is too large
 

- average moisture content is approximately 25 (or 2.5%)
- standard deviation of about 6

**Examine sources of variation:**

**Data Collection:** Hierarchical (or nested) Study Design

- Sample  $b$  barrels of pigment paste
- $s$  samples are taken from the content of each barrel
- Each sample is mixed and divided into  $r$  parts. Each part is sent to the lab.

There are

$$n = (b)(s)(r) \text{ observations.}$$

**Model:**

$$y_{ijk} = \mu + \beta_i + \delta_{ij} + e_{ijk}$$

where

- $y_{ijk}$  is the moisture content determination for the  $k$ -th part of the  $j$ -th sample from the  $i$ -th barrel
- $\mu$  is the mean moisture content
- $\beta_i$  is a random barrel effect:  

$$\beta_i \sim NID(0, \sigma_\beta^2)$$
- $\delta_{ij}$  is a random sample effect:  

$$\delta_{ij} \sim NID(0, \sigma_\delta^2)$$
- $e_{ijk}$  corresponds to random measurement error:  

$$e_{ijk} \sim NID(0, \sigma_e^2)$$

**Covariance Structure:**

- Homogeneous variances:  $V(Y_{ijk}) =$
- Two parts of one sample:  $Cov(Y_{ijk}, Y_{ij\ell}) =$
- Observations from different barrels:  $Cov(Y_{ijk}, Y_{cm\ell}) =$

In this study

- $b = 15$  barrels were sampled
- $s = 2$  samples were taken from each barrel
- $r = 2$  sub-samples were analyzed from each sample taken from each barrel

Write this model in the form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where

$$\mathbf{R} =$$

$$\mathbf{M} =$$

Then

$$E(\mathbf{y}) =$$

$$V(\mathbf{y}) =$$

**Example 3.**

A study was conducted on human subjects to measure the effects of 8 different foods on serum glucose levels. On each of 5 days, 8 subjects randomly assigned to the 8 food with one subject for each food. Each subject was given an assigned food, and then blood samples were taken from each subject at 4 time points: 15, 30, 45, and 60 minutes after consumption of the food. The same process was repeated each day with 8 different subjects, so a total of 40 subjects were used in the experiment. The level of the protein of interest was measured in each of the 160 blood samples. For  $i = 1, \dots, 5$ ,  $j = 1, \dots, 8$ , and  $k = 1, \dots, 4$ , let  $y_{ijk}$  be the serum glucose measurement on day  $i$  for food  $j$  at time  $k$ . For  $i = 1, \dots, 5$ ,  $j = 1, \dots, 8$ , and  $k = 1, \dots, 4$ , consider the model

$$y_{ijk} = \mu_{jk} + d_i + e_{ijk},$$

where  $\mu_{jk}$  terms are unknown fixed parameters and the other terms are random effects defined as follows. Let  $\mathbf{d} = [d_1, \dots, d_5]^T$ . For  $i = 1, \dots, 5$  and  $j = 1, \dots, 8$ , let  $\mathbf{e}_{ij} = [e_{ij1}, \dots, e_{ij4}]^T$ . Suppose

$$\mathbf{d} \sim N(\mathbf{0}, \sigma_d^2 \mathbf{I}_{5 \times 5}),$$

and

$$\mathbf{e}_{ij} \sim N(\mathbf{0}, \Sigma_e) \text{ for } i = 1, \dots, 5 \text{ and } j = 1, \dots, 8,$$

where  $\sigma_d^2$  is an unknown positive variance parameter and

$$\Sigma_e = \sigma_e^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

for some unknown variance parameter  $\sigma_e^2 > 0$  and some unknown correlation parameter  $\phi$ . Finally, suppose that  $\mathbf{d}$  and  $\mathbf{e}_{11}, \dots, \mathbf{e}_{58}$  are all independent. In terms of model parameters, give a simplified expression for the variance of the generalized least squares estimator of each of the following:

$$(a) \mu_{31}$$

(b)  $\bar{\mu}_3$ .

## 6.2 Analysis of Mixed Linear Models

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where  $\mathbf{X}_{n \times p}$  and  $\mathbf{Z}_{n \times q}$  are known model matrices and

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}\right)$$

Then

$$Y \sim N(\mathbf{X}\boldsymbol{\beta}, \Sigma)$$

where

$$\Sigma = \mathbf{Z}\mathbf{M}\mathbf{Z}^T + \mathbf{R}$$

### Some objectives

- (i) Inferences about estimable functions of fixed effects:  
Point estimates, Confidence intervals and Tests of hypotheses.
- (ii) Estimation of variance components (elements of  $\mathbf{M}$  and  $\mathbf{R}$ )
- (iii) Predictions of random effects (blup)
- (iv) Predictions of future observations

### 6.3 Generalized Least Squares (GLS) Estimation

Suppose

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$$

and also suppose

$$\boldsymbol{\Sigma} = V(\mathbf{y}) = \mathbf{Z}\mathbf{M}\mathbf{Z}^T + \mathbf{R}$$

is known. Then a GLS estimator for  $\boldsymbol{\beta}$  is any  $\mathbf{b}$  that minimizes

$$Q(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\mathbf{b})$$

The estimating equations are:

$$(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})\mathbf{b} = \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$$

and

$$\mathbf{b}_{GLS} = (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y})$$

is a solution. For any estimable function  $\mathbf{C}^T \boldsymbol{\beta}$ , the unique b.l.u.e. is

$$\mathbf{C}^T \mathbf{b}_{GLS} = \mathbf{C}^T (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$$

with

$$V(\mathbf{C}^T \mathbf{b}_{GLS}) = \mathbf{C}^T (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{C}$$

If  $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$ , then

## 6.4 Variance component estimation

- Estimation of parameters in  $\mathbf{M}$  and  $\mathbf{R}$
- Crucial to the estimation of estimable functions of fixed effects (e.g.  $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ )
- Of interest in its own right (sources of variation in the pigment paste production example)

### 6.4.1 Rules for Expected Mean Squares

- **Rule 1.** The error term in the model,  $e_{(ij\dots)m}$ , where the subscript  $m$  denotes the replication subscript.
- **Rule 2.** In addition to an  $\mu$  and  $e_{(ij\dots)m}$ , the model contains all the main effects and any interactions that the experimenter assumes exist.
- **Rule 3.** For each term in the model, divide the subscripts into three classes:
  1. live-those subscripts that are present in the term and are not in the parentheses;
  2. dead-those subscripts that are present in the term and are in parentheses; and
  3. absent-those subscripts that are present in the model but not in that particular term.

Thus, in  $(\alpha\beta)_{ij}$ ,  $i$  and  $j$  are live and  $k$  is absent, and in  $e_{(ij)k}$ ,  $k$  is live and  $i$  and  $j$  are dead.

- **Rule 4.** Degrees of freedom. The number of DF for any term in the model is the product of the number of levels associated with each dead subscript and the number of levels minus 1 associated with each live subscript. For example, the number of DF associated with  $(\alpha\beta)_{ij}$  is  $(a-1)(b-1)$ , and the number of DF associated with  $e_{(ij)k}$  is  $ab(n-1)$ .

- **Rule 5.** Each term in the model has either a variance component (random effect) or a fixed factor (fixed effect) associated with it. If an interaction contains at least one random effect, the entire interaction is considered random. A variance component has Greek letters as subscripts to identify the particular random effect. Thus, in a two-factor mixed model with factor  $A$  fixed and factor  $B$  random, the variance component for  $B$  is  $\sigma_b^2$ , and the variance component for  $AB$  is  $\sigma_{ab}^2$ . A fixed effect is always represented by the sum of squares of the model components associated with the factor divided by its degrees of freedom. For example, the effect of  $A$  is

$$\frac{\sum_{i=1}^a \alpha_i^2}{a-1}$$

- **Rule 6.** Expected mean squares. To obtain EMS, prepare the following table. There is a row for each model component (mean square) and a column for each subscript. Over each subscript, write the number of levels of the factor associated with that subscript and whether the factor is fixed (F) or random (R). Replicates are always considered to be random.

1. In each row, write 1 if one of the dead subscripts in the row component matches the subscript in the column:

Factor	F a i	F b j	R n k
$\alpha_i$			
$\beta_j$			
$(\alpha\beta)_{ij}$	1		
$e_{(ij)k}$		1	

- 2. In each row, if any of the subscripts on the row component match the subscript in the column, write 0 if the column is headed by a fixed factor and 1 if the column is headed by a random factor:

Factor	F a i	F b j	R n k
$\alpha_i$	0		
$\beta_j$		0	
$(\alpha\beta)_{ij}$	0	0	
$e_{(ij)k}$	1	1	1

3. In the remaining empty row positions, write the number of levels shown above the column heading:

Factor	F a i	F b j	R n k
$\alpha_i$	0		
$\beta_j$		b	n
$(\alpha\beta)_{ij}$	0	0	n
$e_{(ij)k}$	1	1	1

4. To obtain the EMS for any model component. First cover all column headed by live subscripts on that component. Then, in each row that contains at least the same subscripts as those on the component being considered, take the product of the visible numbers and multiply by the appropriate fixed or random factor from rule 5. The sum of these quantities is the EMS of the model component being considered.

	$F$	$F$	$R$	
Factor	$a$	$b$	$n$	
	$i$	$j$	$k$	
$\alpha_i$	0	$b$	$n$	$\sigma^2 + bn \sum \alpha_i^2 / (a - 1)$
$\beta_j$	$a$	0	$n$	$\sigma^2 + an \sum \beta_i^2 / (b - 1)$
$(\alpha\beta)_{ij}$	0	0	$n$	$\sigma^2 + n \sum (\alpha\beta)_{ij}^2 / (a-1)(b-1)$
$e_{(ij)k}$	1	1	1	$\sigma^2$

**Example 4.**

Use the rules for expected mean squares to derive the expected mean squares for the following models, and propose appropriate test statistics for all effects:

$$(a) y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$$

$$i = 1, 2, \dots, a$$

$$j = 1, 2, \dots, b$$

$$k = 1, 2, \dots, n$$

Assuming that all the factors are random.

$$(b) y_{ijk} = \mu + \alpha_i + b_j + (ab)_{ij} + e_{ijk}$$

$$i = 1, 2, \dots, a$$

$$j = 1, 2, \dots, b$$

$$k = 1, 2, \dots, n$$

Assuming that factor  $A$  is fixed and factor  $B$  is random.

### 6.4.2 ANOVA method (Method of Moments)

- Compute an ANOVA table
- Equate mean squares to their expected values
- Solve the resulting equations

**Example 5.** Penicillin production

$$Y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

where  $\beta_j \sim NID(0, \sigma_\beta^2)$  and  $e_{ij} \sim NID(0, \sigma_e^2)$

Source of Variation	d.f.	Sums of Squares
Blocks	4	$a \sum_{j=1}^b (\bar{Y}_{..j} - \bar{Y}_{..})^2 = SS_{blocks}$
Processes	3	$b \sum_{i=1}^a (\bar{Y}_{i..} - \bar{Y}_{..})^2 = SS_{processes}$
error	12	$\sum_{i=1}^a \sum_{j=1}^b (Y_{ij} - \bar{Y}_{i..} - \bar{Y}_{..j} + \bar{Y}_{..})^2 = SSE$
C. total	19	$\sum_{i=1}^a \sum_{j=1}^b (Y_{ij} - \bar{Y}_{..})^2$

Start at the bottom:

$$MS_{error} = \frac{SSE}{(a-1)(b-1)}$$

$$E(MS_{error}) = \sigma_e^2$$

Then an unbiased estimator for  $\sigma_e$  is

$$\hat{\sigma}_e^2 = MS_{error}$$

Next, consider the mean square for the random block effects:

$$MS_{blocks} = \frac{SS_{blocks}}{b-1}$$

$$E(MS_{blocks}) = \sigma_e^2 + a\sigma_\beta^2$$

Then,

$$\begin{aligned}\sigma_\beta^2 &= \frac{E(MS_{blocks}) - \sigma_e^2}{a} \\ &= \frac{E(MS_{blocks}) - E(MS_{error})}{a} \\ \hat{\sigma}_\beta^2 &= \frac{MS_{blocks} - MS_{error}}{a}\end{aligned}$$

An unbiased estimator for  $\sigma_\beta^2$  is

For the penicillin data

$$\hat{\sigma}_e^2 =$$

$$\hat{\sigma}_\beta =$$

$$\widehat{V(Y_{ij})} =$$

### 6.4.3 Properties of ANOVA methods for variance component estimation

- (i) Broad applicability
  - easy to compute in balanced cases
  - ANOVA is widely known
  - not required to completely specify distributions for random effects
- (ii) Unbiased estimators
- (iii) Sampling distribution is not exactly known, even under the usual normality assumptions (except for  $\hat{\sigma}_e^2 = MS_{error}$ )
- (iv) May produce negative estimates of variances
- (vi) For unbalanced studies, there may be no “natural” way to choose

**Result 1.** If  $MS_1, MS_2, \dots, MS_k$  are distributed independently with

$$\frac{(df_i)MS_i}{E(MS_i)} \sim \chi_{df_i}^2$$

and constants  $a_i > 0$ ,  $i = 1, 2, \dots, k$  are selected so that

$$\hat{\sigma}^2 = \sum_{i=1}^k a_i MS_i$$

has expectation  $\sigma^2$ , then

$$V(\hat{\sigma}^2) = 2 \sum_{i=1}^k \frac{a_i^2 [E(MS_i)]^2}{df_i}$$

and an unbiased estimator of this variance is

$$\hat{V}(\hat{\sigma}^2) = \frac{2 \sum a_i^2 MS_i^2}{(df_i + 2)}$$

(v) For unbalanced studies, there may be no “natural” way to choose

$$\hat{\sigma}^2 = \sum_{i=1}^k a_i MS_i$$

Furthermore,

$$E(MS_i^2) = V(MS_i) + [E(MS_i)]^2$$

$$\begin{aligned} &= \frac{2[E(MS_i)]^2}{df_i} + [E(MS_i)]^2 \\ &= \left( \frac{df_i + 2}{df_i} \right) [E(MS_i)]^2 \end{aligned}$$

Consequently,

$$E \left[ 2 \sum_{i=1}^k \frac{a_i^2 MS_i^2}{(df_i + 2)} \right] = V(\hat{\sigma}^2)$$

A “standard error” for

$$\hat{\sigma}^2 = \sum_{i=1}^k a_i MS_i$$

could be reported as

$$S_{\hat{\sigma}^2} = \sqrt{\frac{2 \sum_{i=1}^k \frac{a_i^2 MS_i^2}{(df_i + 2)}}{}}$$

Using the Cochran-Satterthwaite approximation, an approximate  $(1 - \alpha) \times 100\%$  confidence interval for  $\sigma^2$  could be constructed as:

$$\begin{aligned} 1 - \alpha &\doteq Pr \left\{ \chi_{\nu,1-\alpha/2}^2 \leq \frac{v\hat{\sigma}^2}{\sigma^2} \leq \chi_{\nu,\alpha/2}^2 \right\} \\ &= Pr \left\{ \frac{v\hat{\sigma}^2}{\chi_{\nu,\alpha/2}^2} \leq \sigma^2 \leq \frac{v\hat{\sigma}^2}{\chi_{\nu,1-\alpha/2}^2} \right\} \end{aligned}$$

where  $\hat{\sigma}^2 = \sum_{i=1}^k a_i MS_i$  and

$$v = \frac{\left[ \sum_{i=1}^k a_i MS_i \right]^2}{\sum_{i=1}^k \frac{[a_i MS_i]^2}{df_i}}$$

**Example 6.** Pigment production

In this example the main objective is the estimation of the variance components

Source of Variation	d.f.	MS	E(MS)
Batches	15-1=14	86.495	$\sigma_e^2 + 2\sigma_\delta^2 + 4\sigma_\beta^2$
Samples in Batches	15(2-1)=15	57.983	$\sigma_e^2 + 2\sigma_\delta^2$
Tests in Samples	(30)(2-1)=30	0.917	$\sigma_e^2$

(a) Estimates of variance components.

- (b) Find a 95% confidence interval for  $\mu$ .

**Example 7.**

A study was conducted on human subjects to measure the effects of 3 foods on serum glucose levels. Each of the 3 foods was randomly assigned to 6 subjects. The serum glucose was measured for each of the subjects at 6 different time points starting at 15 minutes and every 15 minutes after food was ingested. Consider the model

$$Y_{ijk} = \mu + \alpha_i + S_{ij} + \tau_k + \gamma_{ik} + e_{ijk}$$

where  $y_{ijk}$  is the serum glucose levels at the  $k^{th}$  time point for the  $j^{th}$  subject with the  $i^{th}$  food,  $\alpha_i$  is the fixed diet effect,  $\tau_k$  is the fixed time effect, and  $\gamma_{ik}$  is the fixed diet  $\times$  time effect,  $S_{ij} \sim NID(0, \sigma_S^2)$  and is independent of  $e_{ijk} \sim NID(0, \sigma_e^2)$ .

(a) Find  $V(\mathbf{Y}_{ij})$ , for this model?

(b) Provide the formulas for the estimator of  $\sigma_e^2$  and  $\sigma_S^2$ .

(c) What is the correlation between observations taken on the same subject?

(d) Find the estimator of  $V(\bar{Y}_{ij})$  and provide it's degrees of freedom.

(e) Find the estimator of  $V(\bar{Y}_{i,k})$  and provide it's Satterthwaith degrees of freedom.

## 6.5 Estimation of fixed effects

Denote the resulting REML estimators as

$$\hat{\mathbf{M}} \quad \hat{\mathbf{R}} \quad \text{and} \quad \hat{\Sigma} = \mathbf{Z}\hat{\mathbf{M}}\mathbf{Z}^T + \hat{\mathbf{R}}$$

For any estimable function  $\mathbf{C}\boldsymbol{\beta}$ , the blue is the generalized least squares estimator

$$\mathbf{C}\mathbf{b}_{GLS} = \mathbf{C}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{y}$$

Using the REML estimator for

$$\Sigma = \mathbf{Z}\mathbf{M}\mathbf{Z}^T + R$$

an approximation is

$$\mathbf{C}\hat{\boldsymbol{\beta}} = \mathbf{C}(\mathbf{X}^T \hat{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\Sigma}^{-1} \mathbf{y}$$

and for “large” samples:

$$\mathbf{C}\hat{\boldsymbol{\beta}} \sim N(\mathbf{C}\boldsymbol{\beta}, \mathbf{C}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{C}^T)$$

## 6.6 Prediction of random effects

Given the observed responses  $\mathbf{y}$ , predict the value of  $\mathbf{u}$ .

For our model,

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{M} & 0 \\ 0 & R \end{bmatrix}\right).$$

Then

$$\begin{aligned} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \end{bmatrix} &= \begin{bmatrix} \mathbf{u} \\ \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} \\ \mathbf{X}\boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{Z} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \\ &\sim N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{X}\boldsymbol{\beta} \end{bmatrix}, \begin{bmatrix} \mathbf{M} & \mathbf{M}\mathbf{Z}^T \\ \mathbf{Z}\mathbf{M} & \mathbf{Z}\mathbf{M}\mathbf{Z}^T + R \end{bmatrix}\right) \end{aligned}$$

The Best Linear Unbiased Predictor (BLUP) is the b.l.u.e. for

$$\begin{aligned} E(\mathbf{u}|\mathbf{y}) &= E(\mathbf{u}) + (\mathbf{M}\mathbf{Z}^T)(\mathbf{Z}\mathbf{M}\mathbf{Z}^T + R)^{-1}(\mathbf{y} - E(\mathbf{y})) \\ &= \mathbf{0} + \mathbf{M}\mathbf{Z}^T(\mathbf{Z}\mathbf{M}\mathbf{Z}^T + R)^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned}$$

Substitute the b.l.u.e. for  $\mathbf{X}\boldsymbol{\beta}$ ,

$$\begin{aligned} \mathbf{X}\mathbf{b}_{GLS} &= \mathbf{X}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{y} \\ \text{Then, the BLUP for } \mathbf{u} \text{ is} \\ BLUP(\mathbf{u}) &= \mathbf{M}\mathbf{Z}^T \Sigma^{-1}(\mathbf{y} - \mathbf{X}\mathbf{b}_{GLS}) \\ &= \mathbf{M}\mathbf{Z}^T \Sigma^{-1}(\mathbf{I} - \mathbf{X}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1})\mathbf{y} \end{aligned}$$

When  $\mathbf{M}$  and  $\boldsymbol{\Sigma} = \mathbf{Z}\mathbf{M}\mathbf{Z}^T + R$  are known. Substituting REML estimators  $\hat{\mathbf{M}}$  and  $\hat{\mathbf{R}}$  for  $\mathbf{M}$  and  $\mathbf{R}$ , an approximate BLUP for  $\mathbf{u}$  is

$$\begin{aligned}\hat{\mathbf{u}} &= \hat{\mathbf{M}}\mathbf{Z}^T\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{I} - \mathbf{X}(\mathbf{X}^T\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\hat{\boldsymbol{\Sigma}}^{-1})\mathbf{y} \\ &= \hat{\mathbf{M}}\mathbf{Z}^T\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y} - \underline{\mathbf{X}\hat{\beta}})\end{aligned}$$

For “large” samples, the distribution of  $\hat{\mathbf{u}}$  is approximately multivariate normal with mean vector  $\mathbf{0}$  and covariance matrix

$$\mathbf{M}\mathbf{Z}^T\boldsymbol{\Sigma}^{-1}(\mathbf{I} - P)\boldsymbol{\Sigma}(\mathbf{I} - P)\boldsymbol{\Sigma}^{-1}\mathbf{Z}\mathbf{M}$$

where

$$P = \mathbf{X}(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}$$

Given estimates  $\hat{\mathbf{M}}$ ,  $\hat{\mathbf{R}}$  and  $\hat{\boldsymbol{\Sigma}} = \mathbf{Z}\hat{\mathbf{M}}\mathbf{Z}^T + \hat{\mathbf{R}}$ ,

$\hat{\beta}$  and  $\hat{\mathbf{u}}$  provide a solution to the mixed model equations:

$$\begin{bmatrix} \mathbf{X}^T\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}^T\hat{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}^T\hat{\mathbf{R}}^{-1} & \mathbf{Z}^T\hat{\mathbf{R}}^{-1}\mathbf{Z} + \hat{\mathbf{M}}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T\hat{\mathbf{R}}^{-1}\mathbf{y} \\ \mathbf{Z}^T\hat{\mathbf{R}}^{-1}\mathbf{y} \end{bmatrix}$$

A generalized inverse of

$$\begin{bmatrix} \mathbf{X}^T\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}^T\hat{\mathbf{R}}^{-1}\mathbf{Z} \\ \mathbf{Z}^T\hat{\mathbf{R}}^{-1} & \mathbf{Z}^T\hat{\mathbf{R}}^{-1}\mathbf{Z} + \hat{\mathbf{M}}^{-1} \end{bmatrix}$$

is used to approximate the covariance matrix for  $\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix}$